DOCUMENT RESUME

ED 053 172 TM 000 692

AUTHOR Aiken, Lewis R., Jr.
TITLE Interactions Among Group Regressions: An Old Method in a New Setting.
INSTITUTION Stanford Univ., Calif. Stanford Center for Research and Development in Teaching.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
REPORT NO RDM-42
PUB DATE Dec 68
CONTRACT OEC-6-10-078
NOTE 16p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Educational Research, *Interaction, Mathematics, *Multiple Regression Analysis, *Prediction, Psychological Studies, *Scores, Tests of Significance
IDENTIFIERS *Neyman Johnson Technique

ABSTRACT

The purpose of the Neyman-Johnson statistical technique is to determine a region or span of values on r independent variables where the predicted criterion scores of two or more treatment groups are significantly different. Consequently, the technique should prove especially useful in research concerned with moderator variables or with the interactions between treatments and person variables. The mechanics of the technique are reviewed and some extensions mentioned. Three simple examples are given. (Author)

SC
RD
T

Stanford Center for Research and Development in Teaching

SCHOOL OF EDUCATION STANFORD UNIVERSITY

STANFORD CENTER
FOR RESEARCH AND DEVELOPMENT
IN TEACHING

Research and Development Memorandum No. 42

INTERACTIONS AMONG GROUP REGRESSIONS: AN OLD METHOD IN A NEW SETTING

Lewis R. Aiken, Jr.

School of Education
Stanford University
Stanford, California

December 1968

This paper was prepared under the provisions of a contract with the United States Department of Health, Education, and Welfare, Office of Education (Contract No. OE-6-10-078, Project No. 5-0252-0501).

PREFACE

The Neyman-Johnson technique is an old method with new importance, as Professor Aiken's title suggests. In this paper the method is explained, with applications to show its usefulness for educational research in general and Center-related activities in particular.

Professor Aiken is visiting the Stanford Center during this academic year as a USOE Postdoctoral Fellow in Educational Research. During the 1969-1970 academic year he will join the faculty of Guilford College, Greensboro, N. C., as Professor of Psychology and Chairman of the Psychology Department.

Richard E. Snow
Coordinator, Program on Heuristic Teaching

Abstract

The purpose of the Neyman-Johnson statistical technique is to determine a region or span of values on r independent variables where the predicted criterion scores of two or more treatment groups are significantly different. Consequently, the technique should prove especially useful in research concerned with moderator variables or with the interactions between treatments and person variables. The mechanics of the technique are reviewed and some extensions mentioned. Three simple examples are given.

Interactions Among Group Regressions: An Old Method in a New Setting

Lewis R. Aiken, Jr.

Although multivariate statistical methods appropriate for the analysis of educational and psychological data are now readily available, some of the potentially most useful methods are unfamiliar to many researchers in education and psychology. One such example is the Neyman-Johnson technique for testing differences among group regressions, a statistical procedure introduced over 30 years ago (Johnson & Neyman, 1936) and extended somewhat during the ensuing years (Abelson, 1953; Potthoff, 1964) but still not commonly known.

To be sure, there are examples in the older literature of studies which have employed the Neyman-Johnson technique (Hansen, 1944; D. A. Johnson, 1949; H. C. Johnson, 1944; Johnson & Fay, 1950; Johnson & Hoyt, 1947), but these papers, written by a few sophisticates, are either insufficiently clear on how the technique was employed or perhaps too replete with complicated symbolism for the majority of readers who might find the technique useful.

In its most general formulation, the Neyman-Johnson technique is a procedure for determining into which of two or more categories (e.g., treatment conditions) an individual with a certain set of scores on r independent (control) variables should be placed in order to maximize his score on a

criterion variable. The problem is a contemporary one, considering the current interest in moderator variables and aptitude-treatment interactions. One purpose of the present paper is to indicate that in these types of investigations there are alternatives to tests for parallelism (common slope) of a set of regression lines.

Formulation and Extensions

In the original formulation of the problem (Johnson & Neyman, 1936), two treatment groups (1 and 2), two independent variables (x_1 and x_2) and one dependent variable (y) were specified. The task is to find the set(s) of points (x_1, x_2), in the space having the two independent variables as axes, in which the y value predicted from the regression equation for group 2 is significantly larger or smaller than the y value predicted from the regression equation for group 1. Such sets of points or regions are specified by a quadratic equation which plots as a conic section (ellipsoid). Abelson (1953), in a rather lucid presentation of the mechanics of the technique, generalized it to three or more independent variables. A further extension by Potthoff (1964), which is somewhat more difficult to follow because of an error in his formula 2.4, considered the procedure when the number of groups is greater than two and the number of criterion variables greater than one. Potthoff also argued for slightly different procedures depending on the research question. Thus, different computations are involved when one wishes to determine whether two treatment conditions are different for a certain point (x_1,

$x_2, \ldots, x_r$) in the region of significance demarcated by the Neyman-Johnson technique, in contrast to determining whether the two treatments are different simultaneously for all points in the region. In addition, Potthoff recommended the construction of confidence limits for the difference between the regression equations of the two groups as a feasible alternative to the plotting of regions of significance, especially when the number of independent variables is greater than two. Finally, Potthoff cautioned that the Neyman-Johnson technique may result in significance regions that are too small or outside the range of actual values on the independent variables, or confidence limits that are too broad to be of use. This is particularly likely when the numbers of cases in the treatment groups are small and/or the residual variances in the regression analyses of the scores of the two groups are large.

Preliminary Tests

Abelson (1953) suggested a list of steps or assumptions that may serve as a guide to when the Neyman-Johnson technique should be used. Given two groups (1 and 2):

1. Determine whether the residual variances (the variances of the observed y's about the regression surface) are significantly different for the two groups.
2. If the residual variances are not significantly different, test for parallelism of the regressions of the two groups, i.e. $(\beta_{11}, \beta_{21}, \ldots, \beta_{r1}) = (\beta_{21}, \beta_{22}, \ldots \beta_{r2})$.
3. If the regressions are not parallel, test for equality of intercepts of the two groups, i.e., $\beta_{01} = \beta_{02}$. If

the regressions are significantly non-parallel, use the Neyman-Johnson technique.

Statistics of the Neyman-Johnson Technique

In order to make the statistical procedure more comprehensible to a wider audience and more consistent with general statistical notation, Abelson's (1953) and Potthoff's (1964) notations have been modified to some extent in the present paper. The majority of the required values can be obtained from a conventional multiple regression program and the additional procedures carried out quite easily on a computer or a desk calculator.

Let subscript i stand for the ith independent variable ($i = 1, 2, \ldots, r$), subscript j the jth group ($j = 1, 2$), and subscript k the kth person ($k = 1, 2, \ldots, n_j$). Then x_{ijk} is the raw score of person k in group j on independent variable i, and y_{jk} is that person's score on the dependent (criterion) variable. The n_j by $(r + 1)$ matrix of scores x_{ijk} and the vector of n_j scores y_{jk} are:

$$X_j = \begin{bmatrix} 1 & x_{1j1} & x_{2j1} & \cdots & x_{rj1} \\ 1 & x_{1j2} & x_{2j2} & \cdots & x_{rj2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 1 & x_{1jn_j} & x_{2jn_j} & \cdots & x_{rjn_j} \end{bmatrix} \qquad \underline{y}_j = \begin{bmatrix} y_{j1} \\ y_{j2} \\ \vdots \\ y_{jn_j} \end{bmatrix}$$

The vector of intercepts and partial regression coefficients for group j is computed as $\underline{b}_j = (X_j'X_j)^{-1}X_j'\underline{y}_j$, and define $(\underline{b}_2' - \underline{b}_1') = (b_{02} - b_{01}, b_{12} - b_{11}, \ldots, b_{r2} - b_{r1})$.

The combined residual sum of squares for the two groups is $ss_e = \sum_{j=1}^{2} (\underline{y}_j'\underline{y}_j - \underline{b}_j'X_j'\underline{y}_j)$, and defining the pooled residual

degrees of freedom as $f = \sum_{j=1}^{2} (n_j - r - 1)$, the mean square for error is $ms_e = ss_e/f$. In order to estimate the variance of the difference between the regression equations for the two groups, compute $V = ms_e[(X_1'X_1)^{-1} + (X_2'X_2)^{-1}]$. Finally, let vector $\underline{x}' = (x_0, x_1, x_2, \ldots, x_r)$ be a list of hypothetical raw scores on the independent variables, where $x_0 = 1$.

Finding Regions of Significance and Confidence Limits

There are two possibilities to consider in setting up a quadratic equation for determining the x region(s) of significance or confidence limits for the differences between predicted y's. To find a critical region such that, with confidence $100(1 - \alpha)$, it can be stated that the two groups are different for any individual point in the region, compute:

(1) $\underline{x}'[(\underline{b}_2 - \underline{b}_1)(\underline{b}_2' - \underline{b}_1') - F_{1,f;\alpha}V]\underline{x} \geq 0.$

On the other hand, to find a critical region such that, with confidence 100(1 −), it can be stated that the two groups are different simultaneously for all points contained in the region, compute:

(2) $\underline{x}'[(\underline{b}_2 - \underline{b}_1)(\underline{b}_2' - \underline{b}_1') - (r + 1)F_{r+1,f;\alpha}V]\underline{x} \geq 0.$

Potthoff suggested that, since plotting regions by use of the above equations is so tedious, the investigator may simply settle for constructing confidence limits for the expression $(\beta_2' - \beta_1')\underline{x}$ as:

(3) $[(\underline{b}_2' - \underline{b}_1')\underline{x}] \pm t_{f;\alpha/2}(\underline{x}'V\underline{x})^{\frac{1}{2}}$, or

(4) $[(\underline{b}_2' - \underline{b}_1')\underline{x}] \pm \sqrt{(r + 1)F_{r+1,f;\alpha}(\underline{x}'V\underline{x})}$,

corresponding to the region formulas given in 1 and 2, respectively. For a given set of scores $\underline{x}$, if formula 3 does

not include 0 then it can be stated with $100(1 - \alpha)$ per cent confidence that the predicted criterion scores corresponding to $\underline{x}$ are significantly different in the two groups. Formula 4 allows the investigator to make a similar statement simultaneously for all points $\underline{x}$ in the critical region of formula 2.

Plotting the Critical Region(s)

The quadratic equations of formulas 1 and 2 above describe conic sections (ellipsoids) and will give two significance regions--one where the predicted value of y in group 2 is larger than in group 1 and the other where the predicted y value in group 1 is larger than in group 2. The boundaries of these regions are not difficult to compute and plot when r, the number of independent variables, is less than three. This can be accomplished most efficiently when $r = 2$ by substituting successive equally spaced values of x_2 into the quadratic equation, setting the equation equal to 0, and applying the quadratic formula $x_1 = (-b \pm \sqrt{b^2 - 4ac})/2a$ to determine the boundary values of x_1 for the given value of x_2.[1]

Elaborations on the Technique

Although Abelson (1953) and Potthoff (1964) do not explicitly mention the fact, the Neyman-Johnson technique can easily be extended downward to cover the case of one independent variable. In this case the computations are much simpler, since

$$(5)\quad \underline{x}'V\underline{x} = ms_e\left\{\sum_{j=1}^{2} (n_j x^2 - 2x\sum_k x_{jk} + \sum_k x^2_{jk})/[n_j\sum_k x^2_{jk} - (\sum_k x_{jk})^2]\right\},$$

and the other needed quantities may be computed from simple linear regression formulas. The significance region(s) in this

[1]Of course, when $r = 1$, simply setting the one-variable quadratic equation equal to zero and finding the roots by the quadratic formula will do the trick.

case, however, will be demarcated by lines parallel to the y axis (see Figures 1 and 2).

As was indicated above, Potthoff (1964) extended the Neyman-Johnson technique to g groups and p criterion variables. These extensions consist of pairwise comparisons of the g groups on the p criteria and require the plotting of $pg(g - 1)/2$ regions, one for each pair. The extensions are straightforward and do not involve computations greatly different from those detailed above (see Appendix).

Examples

Three examples of applying the Neyman-Johnson technique will be given--two where $r = 1$ and one where $r = 2$. In a dissertation study at Stanford University by Mary Lou Koran, student teachers were exposed to one of two kinds of information between microteaching sessions. The 40 students in group 2 were exposed to a film portrayal of the particular teaching skill to be learned, and the 40 students in group 1 read a verbatim text of the sound track of the film. The skill to be learned was the formulation of analytic questions by the teacher during class discussion. The independent variables are the Hidden Figures test from the Kit of Selected Reference Aptitude and Achievement Factors (French, 1963) and a test called Film Memory. The regressions of the criterion (Total Number of Analytic Questions) scores on Hidden Figures scores are illustrated in Figure 1, and the regressions of the criterion on Film Memory are illustrated in Figure 2. The broken vertical lines in the figures demarcate regions of significance

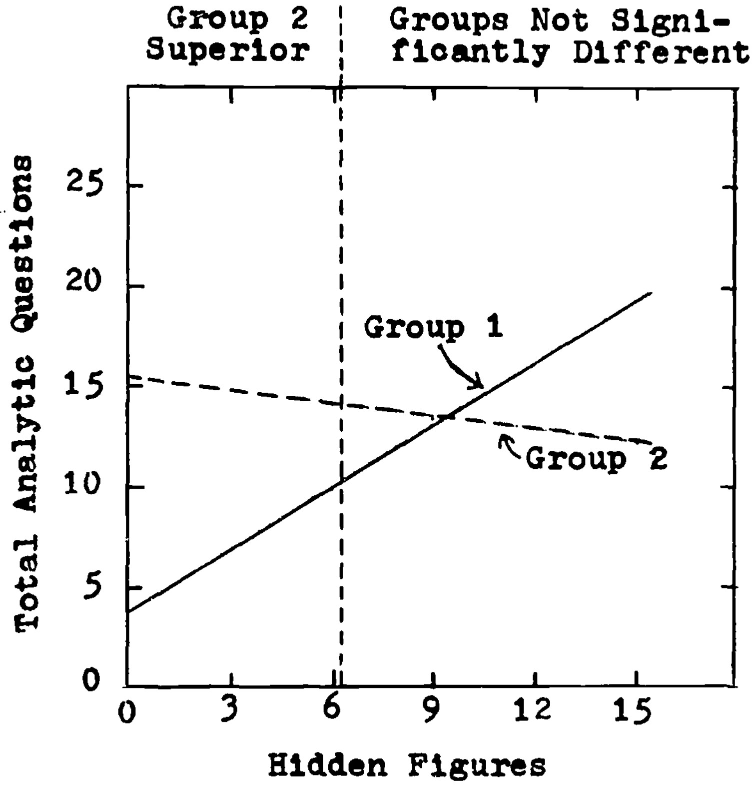

Fig. 1. Regions of significance in example 1 with one independent variable. (See text for explanation.)

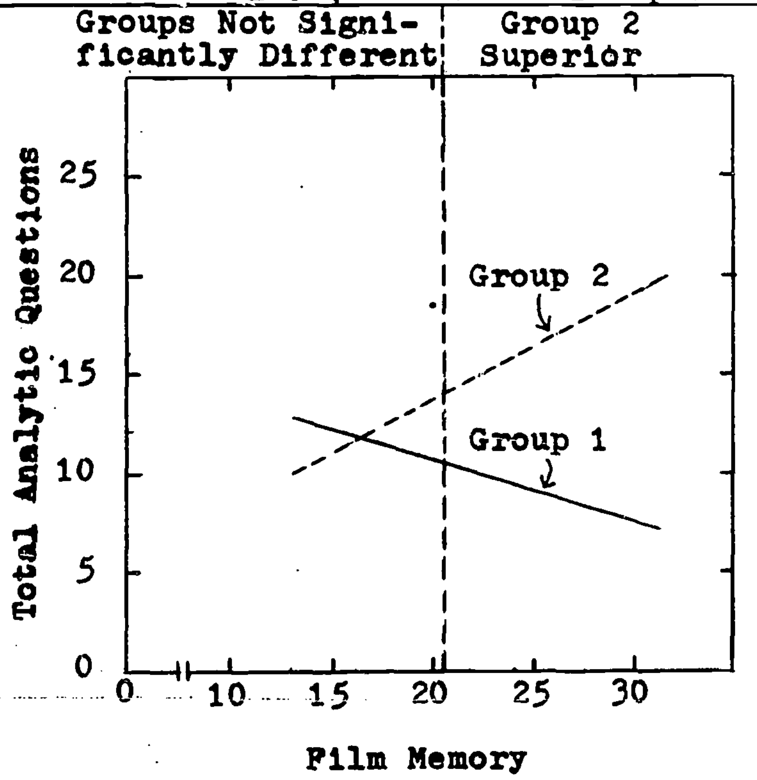

Fig. 2. Regions of significance in example 2 with one independent variable. (See text for explanation.)

(.95 level, formula 2) where the predicted criterion scores of treatment group 2 are significantly different from those of treatment group 1.

In the area to the left of the broken vertical line in Figure 1 (very low scores on Hidden Figures), the predicted criterion scores of group 2 are significantly higher than those of group 1. A similar area where group 1 is superior to group 2 in predicted criterion scores lies to the right of Figure 1, but since this area contains no actual data points it is not shown in the illustration.

In the area to the right of the broken vertical line in Figure 2 (very high scores on Film Memory), the predicted criterion scores of group 2 are significantly higher than those of group 1. A similar area where group 1 is superior to group 2 in predicted criterion scores lies to the left of Figure 2, but since this area contains no actual data points it is not shown in the illustration.

Since it is obvious that the regression slopes are quite different for the two groups in Figures 1 and 2, it was decided to analyze the combined effects of Hidden Figures and Film Memory on Total Analytic Questions. Figure 3 illustrates the solution when both independent variables are considered. The 75% region of significance where group 1 is superior to group 2 is off the graph to the right and contains no data points.[2] Of course, the fact that the predicted criterion scores of one

[2]Since the correlations between the independent variables were essentially zero for both groups in this study, the significance regions in Figure 3 can be roughly predicted from the results depicted in Figures 1 and 2.

group are superior to those of another group in a given region does not necessarily imply that a given treatment should be adopted for all examinees whose x scores fall in that region. Which treatment should be employed with a given individual depends not only on the probability of his scores falling within a certain treatment region but also on such factors as cost and convenience of the treatment.

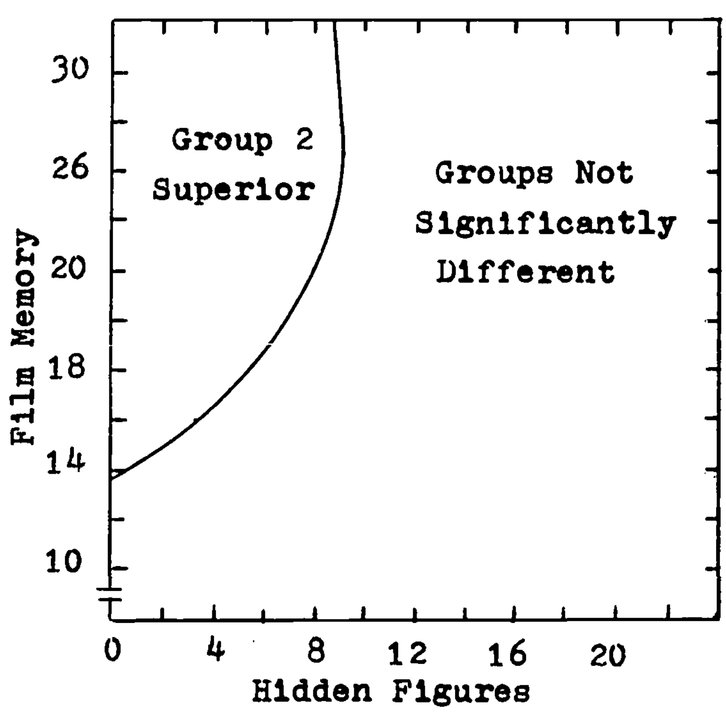

Fig. 3. Regions of significance in example 3 with two independent variables. (See text for explanation.)

References

Abelson, R. P. A note on the Neyman-Johnson technique. Psychometrika, 1953, 18, 213-217.

French, J. W. Kit of selected tests for aptitude and achievement factors. (Rev. ed.) Princeton: Educational Testing Service, 1963.

Hansen, C. W. Factors associated with successful achievement in problem solving in sixth grade arithmetic. *Journal of Educational Research*, 1944, *38*, 111-118.

Johnson, D. A. An experimental study of the effectiveness of films and film strips in teaching geometry. *Journal of Experimental Education*, 1949, *17*, 363-372.

Johnson, H. C. The effect of instruction in mathematical vocabulary upon problem solving in arithmetic. *Journal of Educational Research*, 1944, *38*, 97-110.

Johnson, P. O., & Fay, L. The Neyman-Johnson technique, its theory and applications. *Psychometrika*, 1950, *15*, 349-367.

Johnson, P. O., & Hoyt, C. On determining three-dimensional regions of significance. *Journal of Experimental Education*, 1947, *15*, 342-353.

Johnson, P. O., & Neyman, J. Tests of certain linear hypotheses and their applications to some educational problems. *Statistical Research Memoirs*, 1936, *1*, 57-63.

Potthoff, R. F. On the Johnson-Neyman technique and some extensions thereof. *Psychometrika*, 1964, *29*, 241-256.

Appendix

Neyman-Johnson Technique for More Than Two Groups

Let there be g groups, with (J, j) being any of the $m = g(g-1)/2$ pairs $(J > j)$. In expressions for ss_e and ms_e, $\sum_{j=1}^{2}$ becomes $\sum_{j=1}^{g}$. Formulas for obtaining simultaneous confidence intervals for $(\beta_J' - \beta_j')\underline{x}$ for all possible (J, j) but for a single $\underline{x}$ are:

(6) $(\underline{b}_J' - \underline{b}_j')\underline{x} \pm \sqrt{(g-1)F_{g-1,f;\alpha}\underline{x}'V_{Jj}\underline{x}}$, or

(7) $(\underline{b}_J' - \underline{b}_j')\underline{x} \pm t_{f;\alpha/m}\sqrt{\underline{x}'V_{Jj}\underline{x}}$, whichever is smaller.

To obtain simultaneous confidence intervals for all $\underline{x}$, use:

(8) $(\underline{b}_J' - \underline{b}_j')\underline{x} \pm \sqrt{(r+1)(g-1)F_{(r+1)(g-1),f;\alpha}\underline{x}'V_{Jj}\underline{x}}$, or

(9) $(\underline{b}_J' - \underline{b}_j')\underline{x} \pm \sqrt{(r+1)F_{(r+1),f;\alpha/m}\underline{x}'V_{Jj}\underline{x}}$, whichever is smaller. Formulas for the $m = g(g-1)/2$ regions corresponding to the confidence intervals described above are, for formula 9 for example:

(10) $\underline{x}'[(\underline{b}_J - \underline{b}_j)(\underline{b}_J' - \underline{b}_j') - (r+1)F_{(r+1),f;\alpha/m}V_{Jj}]\underline{x} \geq 0.$

Formulas for regions corresponding to the confidence limits of formulas 6, 7, and 8 may be written in similar fashion.

More Than One Criterion Variable

The problem is to obtain simultaneous confidence limits for the differences $(\underline{\beta}_J^{l}{}' - \underline{\beta}_j^{l}{}')\underline{x}$, where l is the "lth" criterion variable (l = 1, 2, . . . , p). An appropriate formula is:

(11) $(\underline{b}_J^{l}{}' - \underline{b}_j^{l}{}')\underline{x} \pm \sqrt{(r+1)(g-1)F_{(r+1)(g-1),f,\alpha/p}\underline{x}'V_{Jj}\underline{x}}$, or

(12) $(\underline{b}_J^{l}{}' - \underline{b}_j^{l}{}')\underline{x} \pm \sqrt{(r+1)F_{(r+1),f;\alpha/mp}\underline{x}'V_{Jj}\underline{x}}$, whichever is smaller; $\underline{b}_J^l - \underline{b}_j^l$ and ms_e^l are the same functions of the y_{jk}^l as $\underline{b}_J - \underline{b}_j$ and ms_e are, respectively, of the y_{jk}'s. A region approach is equivalent to the above, but there will be p times as many regions as in the univariate case of one y variable.